

FORM PTO-1390 (REV 10-96)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER GDC-129	
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371				U.S. APPLICATION NO. (if known, see 37 CFR 1.5)	
				09/744323	
INTERNATIONAL APPLICATION NO. PCT/US99/16810		INTERNATIONAL FILING DATE 22 July 1999		PRIORITY DATE CLAIMED 22 July 1998	
TITLE OF INVENTION Explicit Rate ABR Algorithm for Use in an ATM Switch					
APPLICANT(S) FOR DO/EO/US HUSSAIN, Iftexhar and BAINS, Kuldip					
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:					
<ol style="list-style-type: none"><li>1. <input checked="" type="checkbox"/> This is a <b>FIRST</b> submission of items concerning a filing under 35 U.S.C. 371.</li><li>2. <input type="checkbox"/> This is a <b>SECOND</b> or <b>SUBSEQUENT</b> submission of items concerning a filing under 35 U.S.C. 371.</li><li>3. <input checked="" type="checkbox"/> This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).</li><li>4. <input checked="" type="checkbox"/> A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.</li><li>5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2))<ol style="list-style-type: none"><li>a. <input checked="" type="checkbox"/> is transmitted herewith (required only if not transmitted by the International Bureau).</li><li>b. <input type="checkbox"/> has been transmitted by the International Bureau.</li><li>c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).</li></ol></li><li>6. <input type="checkbox"/> A translation of the International Application into English (35 U.S.C. 371(c)(2)).</li><li>7. <input type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))<ol style="list-style-type: none"><li>a. <input type="checkbox"/> are transmitted herewith (required only if not transmitted by the International Bureau).</li><li>b. <input type="checkbox"/> have been transmitted by the International Bureau.</li><li>c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.</li><li>d. <input type="checkbox"/> have not been made and will not be made.</li></ol></li><li>8. <input type="checkbox"/> A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).</li><li>9. <input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).</li><li>10. <input type="checkbox"/> A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).</li></ol>					
Items 11. to 16. below concern document(s) or information included:					
<ol style="list-style-type: none"><li>11. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.</li><li>12. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.</li><li>13. <input type="checkbox"/> A FIRST preliminary amendment. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment.</li><li>14. <input type="checkbox"/> A substitute specification.</li><li>15. <input type="checkbox"/> A change of power of attorney and/or address letter.</li><li>16. <input type="checkbox"/> Other items or information:</li></ol>					

3507 R304 PCT/PTO

19 JAN 2001

U.S. APPLICATION NO. <b>09/744323</b> INTERNATIONAL APPLICATION NO. <b>PCT/US99/16810</b>		ATTORNEY'S DOCKET NUMBER <b>GDC-129</b>							
17. <input checked="" type="checkbox"/> The following fees are submitted: <b>BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):</b> Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO ..... <b>\$1000</b> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO ..... <b>\$860</b> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO ..... <b>\$710</b> International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4) ..... <b>\$690</b> International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4) ..... <b>\$100</b> <b>ENTER APPROPRIATE BASIC FEE AMOUNT =</b>		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th colspan="2" style="text-align: left;">CALCULATIONS</th> <th style="text-align: left;">PTO USE ONLY</th> </tr> <tr> <td colspan="2" style="height: 100px;"></td> <td></td> </tr> </table>		CALCULATIONS		PTO USE ONLY			
CALCULATIONS		PTO USE ONLY							
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(e)).		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; text-align: right;">\$ 100.00</td> <td style="width: 50%;"></td> </tr> <tr> <td style="text-align: right;">\$</td> <td></td> </tr> </table>		\$ 100.00		\$			
\$ 100.00									
\$									
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE						
Total claims	28 - 20 =	8	X \$18.00						
Independent claims	4 - 3 =	1	X \$80						
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ \$270						
<b>TOTAL OF ABOVE CALCULATIONS =</b>			<b>\$</b>						
Reduction of 1/2 for filing by small entity, if applicable. A Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28).									
<b>SUBTOTAL =</b>			<b>\$ 324.00</b>						
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f)).									
<b>TOTAL NATIONAL FEE =</b>			<b>\$ 324.00</b>						
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property . +									
<b>TOTAL FEES ENCLOSED =</b>			<b>\$</b>						
Amount to be:			\$						
refunded			\$						
charged			\$						

a. ☒ A check in the amount of \$\_\_\_\_\_ to cover the above fees is enclosed.

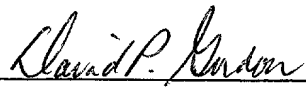
b. ☐ Please charge my Deposit Account No. \_\_\_\_\_ in the amount of \$\_\_\_\_\_ to cover the above fees.  
 A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any  
 overpayment to Deposit Account No. 07-1732. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

**David P. Gordon**  
**65 Woods End Road**  
**Stamford, CT 06905**

  
 SIGNATURE:  
**David P. Gordon**  
 NAME  
**29,996**  
 REGISTRATION NUMBER

EXPLICIT RATE ABR ALGORITHM FOR USE IN AN ATM SWITCH

This application claims the benefit of provisional application serial number 60/093,826 filed July 22, 1998.

## BACKGROUND OF THE INVENTION

## 1. Field of the Invention

The present invention relates broadly to the field of telecommunications. More particularly, the present invention relates to congestion control of available bit rate (ABR) service in an asynchronous transfer mode (ATM) using explicit rate (ER) marking and per-VC queuing.

## 2. State of the Art

Perhaps the most awaited, and now fastest growing technology in the field of telecommunications in the 1990's is known as Asynchronous Transfer Mode (ATM) technology. ATM is providing a mechanism for removing performance limitations of local area networks (LANs) and wide area networks (WANs) and providing data transfers at a speed of on the order of gigabits/second. The variable length packets of LAN and WAN data are being replaced with ATM cells which are relatively short, fixed length packets. Because ATM cells can carry voice, video and data across a single backbone network, the ATM technology provides a unitary mechanism for high speed end-to-end telecommunications traffic.

Current ATM service is offered in different categories according to a user's needs. Some of these categories include constant bit rate (CBR), variable bit rate (VBR), unspecified bit rate (UBR), and available bit rate (ABR). The ABR service category is intended for applications with vague bandwidth and delay requirements. Although the ABR service category potentially can be used for a wide variety of applications, it is primarily intended for data applications requiring low cell loss but tolerant of low bit rates. An ABR connection only

specifies a minimum cell rate (MCR) which may be zero. The ABR service guarantees a low cell loss rate to those ABR connections which adapt their rate in accordance with the feedback from the network.

From the network perspective, the objective of ABR service is to maximize resource utilization by allowing ABR sources to adapt their transmission rates according to the availability of unused resources. In other words, as more bandwidth becomes available, ABR sources are allowed to increase their bit rates. Conversely, when network congestion occurs, ABR sources are directed to decrease their bit rate. The process of adjusting ABR bit rates is generally referred to as congestion control.

There are two common approaches for congestion control of ABR sources: rate based and credit based. The ATM Forum has selected the rate based scheme as the standard for the ABR service. Rate based schemes use feedback information from the network to control the rate at which cells are emitted by an ABR source to the network. The feedback information is carried in special cells called Resource Management (RM) cells. A typical network implementing a feedback control scheme consists of source end systems (SES), destination end systems (DES), and switches. The ATM Forum has specified that SES, DES, and switches must implement at least one of the following methods for congestion control: explicit forward congestion indication (EFCI) marking, relative rate (RR) marking, or explicit rate (ER) marking.

EFCI marking requires that the switch set an EFCI bit in the data cell headers. This marking scheme relies on the DES to convey congestion information to the SES by marking a congestion indication (CI) field in the backward RM cells. RR marking requires that the switch set either the CI field or a no increase (NI) field in the RM cells to convey congestion information. Both EFCI and RR marking are referred to as binary rate feedback. ER marking requires that the switch write the

appropriate rate value in an ER field of the RM cells to specify an explicit rate at which the SES is allowed to transmit data. ER marking is referred to as explicit rate feedback. The transmission rate of an ABR source set by the SES based on the ER feedback is called the allowed cell rate (ACR).

ABR flow control is applied between an SES and a DES which are connected via bidirectional connections. For the sake of simplicity, in the following discussion, only data flow from SES to DES is described. A general end-to-end control loop mechanism for ABR service is shown in prior art Figure 1 where an SES 10 is linked to a DES 12 via several ATM switches 14, 16, 18. The SES generates "forward" RM cells along with the data cells. The DES turns around the forward RM cells and sends them back to the source as "backward" RM cells. The backward RM cells convey feedback information provided by the switches and/or DES. A switch can insert feedback control information into backward or forward RM cells. As described above, a switch must implement at least one of the congestion control methods (EFCI, RR, or ER). The ER field in the forward RM cell is set by the SES to a requested rate such as peak cell rate (PCR). Intermediate switches can only decrease the rate in the ER field in forward and/or backward RM cells.

The ATM Forum has not standardized a specific algorithm for the implementation of congestion control, but has provided five examples in ATM Traffic Management Specification Version 4.0, April 1996, af-tm-0056.000, the complete disclosure of which is hereby incorporated by reference herein. As a result of tremendous research activity regarding ABR, many algorithms for congestion control have been proposed in the literature. The common goal of all of the algorithms is to provide an accurate fair share of available bandwidth for each ABR connection while maximizing link utilization and controlling queue size.

A majority of the known algorithms for ABR congestion control are designed for ATM switches having a common first-in-

first-out ABR queue. Although many of the current ATM switches employ this type of simple queuing and scheduling, the newer generation of ATM switches utilize sophisticated per-VC queuing and scheduling algorithms.

A commonly accepted formula for determining fair share  $F_s$  for each ABR VC at a given link is shown in equation (1) where  $C$  is the total link bandwidth available to the ABR VCs using the link,  $C_b$  is the aggregate rate for the bottlenecked connections,  $N$  is the number of active connections, and  $N_b$  is the number of bottlenecked VCs.

$$F_s = \frac{C - C_b}{N - N_b} \quad (1)$$

A bottleneck connection is a connection which cannot get its fair share at this link because either it is limited by its PCR, or is bottlenecked somewhere else.

Some known switch algorithms attempt to compute an exact fair share using equation (1). Typically, such algorithms need to keep track of per-VC state information (e.g., connection bottleneck status). In addition to the high complexity resulting from the local VC state information, the computed fair share may not be equal to the true fair share because a number of factors introduce inaccuracies in the computation. See, e.g., F. M. Chiussi, A. Arulambalam, Y. Xia, and X. Chen, "Explicit rate ABR schemes using traffic load as congestion indicator," Proc. 6th Int. Conf: Computer Commun. and Networks, Las Vegas, Nevada, Sept. 1997.

#### SUMMARY OF THE INVENTION

It is therefore an object of the invention to provide an explicit rate ABR algorithm for use in an ATM switch.

It is also an object of the invention to provide an explicit rate ABR algorithm which avoids the complexity of known algorithms.

It is another object of the invention to provide an explicit rate ABR algorithm which provides accurate per-VC bandwidth fairness.

It is still another object of the invention to provide an explicit rate ABR algorithm which maximizes link utilization and stable queue control.

In accord with these objects which will be discussed in detail below, the algorithm of the present invention does not attempt to compute the exact fair share, but uses an approximation of the fair share. It relies on the underlying per-VC queuing and scheduling for flow isolation and fair service to the VCs. This yields an efficient algorithm of low complexity that provides per-VC bandwidth fairness, high link utilization, and excellent queue control. The main steps of the algorithm include: determining the available bandwidth, determining the per-VC fair share of available bandwidth, determining the explicit rate, and updating the explicit rate value in the RM cell. The algorithm utilizes an egress side scheduler having a non-work-conserving part (e.g., a traffic shaper) and a round-robin part. The non-work-conserving part satisfies the guaranteed traffic and the round-robin part serves the "best efforts" traffic which may or may not include UBR and VBR-nrt traffic as well as ABR traffic. Priority is given to the non-work-conserving part such that the round-robin part only operates only when the guaranteed traffic has been satisfied. According to the invention, the available bandwidth is determined by measuring the number of cells served by the non-work-conserving part during a measurement interval which is based on the link rate. The per-VC fair share of available bandwidth is calculated in two different ways depending on whether the VC is considered congested. Where the VC is not

considered to be congested, the fair share is calculated by dividing the available bandwidth by an averaged number of currently active best efforts connections served by the round-robin part. Where the VC is considered to be congested, the fair share is calculated by dividing the available bandwidth by the number of currently established best efforts connections and reducing the fair share by a damping factor which is a function of the VC's queue length. A VC is considered active if its queue length is not zero and is considered congested if its queue length exceeds a threshold. Short term fluctuations in the number of active VCs are filtered out through the use of a weighted average. A new explicit rate is determined by adding the fair share to the MCR. The new explicit rate replaces the old explicit rate only if it is lower than the old explicit rate.

Additional objects and advantages of the invention will become apparent to those skilled in the art upon reference to the detailed description taken in conjunction with the provided figures.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a diagram of an SES coupled to a DES via three ATM switches according to the prior art;

Figure 2 is a simplified flow chart illustrating the algorithm of the invention; and

Figure 3 is a diagram of a network configuration used to test the algorithm of the invention.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Figure 2 illustrates the algorithm of the invention which is preferably performed on each backward RM cell. Those skilled in the art will appreciate that the algorithm may be implemented



in hardware, software, or a combination of hardware and software. Turning now to Figure 2, the algorithm starts at 100 and the available bandwidth is calculated at 102. As mentioned above, the algorithm of the invention utilizes an egress side scheduler having a non-work-conserving part (e.g., a traffic shaper) and a round-robin part. The egress scheduler implies that in any finite time interval, bandwidth available for the round-robin scheduler part is the difference of total link capacity and the rate used by the non-work-conserving scheduler part. To approximate the available bandwidth  $A_{bw}$  for the round-robin scheduler part, bandwidth used by the non-work-conserving scheduler part is first estimated. This involves determining the number of cells  $U$  served by the non-work-conserving scheduler part during an interval of  $W$  cell slots. The approximation of the available bandwidth in cells per second is determined according to equation (2).

$$A_{bw} = \left(1.0 - \frac{U}{W}\right) \cdot \text{LinkRate} \quad (2)$$

In experiments applying the algorithm,  $W$  was fixed at 1024 cell slots for a 149.76 Mb/s (OC-3) link. A large value of the measurement interval  $W$  is desirable in order to provide a larger sample size for estimating  $U$  and in order to reduce computational load on the switch (because a larger measurement interval means fewer computation). A large value of the measurement interval  $W$  is undesirable, however, because larger measurement interval means slower (or less frequent) updating of the available bandwidth. As a design trade-off, a value of 1024 cell slots was deemed appropriate for the measurement interval.

After the  $A_{bw}$  is found at 102, a decision is made at 104 whether the particular VC is congested. This determination is made by comparing the number of cells  $Q_{vc}$  in the VC queue to a threshold number  $Q_{thresh}$ . If the number of cells in the queue exceeds the threshold number, the VC is considered congested.

If the VC is congested, the fair share  $F_{share}$  of available bandwidth  $A_{bw}$  is computed at 106 by dividing the  $A_{bw}$  by the number  $N_{est}$  of established connections served by the round-robin scheduler as shown in equation (3).

$$F_{share} = \frac{A_{bw}}{N_{est}} \quad (3)$$

When the VC is congested, the fair share  $F_{share}$  is reduced by a damping factor  $damp(x)$  which is a linear function of the relative congestion of the VC. More particularly, the damping factor is calculated according to equation (4) where  $x = Q_{vc} - Q_{thresh}$  and  $Q_{max}$  is the maximum queue length for the VC.

$$damp(x) = \text{Max} \left( 0, \left( 1.0 - \frac{x}{Q_{max}} \right) \right) \quad (4)$$

A new explicit rate ER is then calculated at 108 by adding the damped fair share to the minimum cell rate MCR as shown in equation (5).

$$ER_{new} = MCR + F_{share} \cdot damp(x) \quad (5)$$

If it is determined at 104 that the VC is not congested, fair share  $F_{share}$  is determined at 112 by the number  $N_{act}$  of active connections served by the round-robin scheduler rather than the number  $N_{est}$  of established connections. A VC is considered active if its  $Q_{vc} > 0$ . The motivation for using active connections in place of established connections is to maintain high link utilization even when "ON/OFF" sources (e.g. "bursty" sources) are present. Before calculating the fair share at 112, however, short term fluctuations in the number of active connections are first filtered out at 110 through the use of a weighted

averaging function. The weighted averaging function is shown as equation (6) where  $\alpha$  is an averaging factor,  $N_{old}$  is the previous measure of the number of active connections, and  $N_{new}$  is the present measure of the number of active connections.

$$N_{act} = (1 - \alpha)N_{old} + \alpha N_{new} \quad (6)$$

The averaging factor  $\alpha$  is chosen on the basis of experience in a particular switch. In the experiments performed by the inventors, described herein below,  $\alpha=1/16$  worked well. After the number of active connections is averaged at 110 using equation (6), the fair share is computed at 112 using equation (7).

$$F_{share} = \frac{A_{bw}}{N_{act}} \quad (7)$$

After the fair share is computed at 112, the new explicit rate for the VC is computed at 114 using equation (8).

$$ER_{new} = MCR + F_{share} \quad (8)$$

After the new explicit rate is computed either at 108 in the case of a congested VC or at 114 in the case of a non-congested VC, the value in the ER field of the RM cell is compared to the new ER at 116. If the new ER is larger than the current ER, no change is made in the RM cell and the algorithm ends (for the particular RM cell) at 118 with respect to this RM cell. If the new ER is smaller than the current ER, the ER field of the RM cell is written at 120 with the new ER value and the algorithm ends (for the particular RM cell) at 122 with respect to this RM cell.

The algorithm of the present invention was tested using a network configuration known as a parking lot which is often used in simulations. The parking lot configuration, shown in Figure 3, includes four switches 200, 202, 204, 206 each having a source 210, 212, 214, 216 of ABR VCs, and a fifth switch 208. The fourth switch 206 also has a source 217 of VBR/UBR traffic. According to the parking lot configuration, each of the switches feeds the next, thereby creating a bottleneck link 218 between the fourth and fifth switches 206, 208.

Each switch 200, 202, 204, 206, 208 is a non-blocking ATM switch with both input and output queuing. On the ingress, traffic is enqueued in a per class queue (with the exception of CBR and VBR-rt which share a common queue) and scheduled on strict priority basis. On the egress, per-VC queuing and scheduling is used. As far as ABR traffic is concerned, the per-VC scheduler is designed to support MCR plus equal share fairness criterion. The ABR traffic sources 210, 212, 214, 216 are persistent (always have data to transmit). The VBR and UBR traffic is modeled by periodic ON/OFF sources which transmits cells at PCR during ON period and no cells during OFF period. All links have a capacity of 149.76 Mb/s, and the length of inter-switch links and the access links is 1000 Km and 1 Km, respectively. The propagation delays through the links is 5 microseconds/Km.

Tests were conducted with the parking lot network loaded by ABR, VBR, and UBR traffic. Two ABR traffic cases corresponding to 10 and 25 VCs were considered. For the ABR sources, the following parameters were kept constant throughout this study: ICR (initial cell rate) = 5 Mb/s, PCR (peak cell rate) = 149.76 Mb/s, RIF (rate increase factor) = 1/128, RDF (rate decrease factor) = 1/256, W = 1024 cell slots and queue threshold  $Q_{\text{thresh}}$  = 16 cells. The performance characteristics measured include: bandwidth fairness, link utilization, and queue control.

The service categories contending for the round-robin scheduler part available bandwidth included: ABR VCs with rate greater than MCR, UBR VCs with rate greater than zero, and VBR-nrt VCs with the rate greater than guaranteed rate.

#### Experiments 1 and 2

Two experiments were conducted with multiple identical ABR VCs: one experiment with ten identical ABR VCs, and one experiment with twenty-five identical ABR VCs. Each VC had an MCR of 1 Mb/s, ICR of 5 Mb/sec and PCR of 149.76 Mb/s. All VCs began transmitting at time zero with a rate equal to ICR.

The results of the first experiment with ten identical ABR VCs showed that the actual transmission rate of ABR VCs converged to the MCR plus an equal share value of about 14.95 Mb/s. The ACR of the ABR sources also converged to this steady state rate. The per-VC queue length converged to a small value in the steady state. During the transient period before steady state, the explicit rate of VCs with shorter round trip delays increased faster resulting in a larger backlog at their buffers. A very high link utilization (near 99%) was achieved.

The results of the second experiment with twenty-five identical ABR VCs showed that the actual transmission rate of ABR VCs converged to the MCR plus equal share value of about 5.99 Mb/s. The ACR of the ABR sources also converged to this rate. The per-VC queue length converged to a small value in the steady state. A very high link utilization (near 99%) was achieved.

#### Experiments 3 and 4

Two experiments were conducted with multiple ABR VCs with different cell rates, one experiment with ten ABR VCs and one experiment with twenty-five ABR VCs.

The third experiment used four ABR VCs each having an MCR of 5 Mb/s and six ABR VCs each having an MCR of 1 Mb/s. All of the VCs had an ICR of 5 Mb/s and PCR of 149.76 Mb/s. All VCs started transmitting at time zero with a rate equal to the ICR. The results of this experiment were that the MCR plus equal share of the four ABR VCs having an MCR of 5 Mb/s converged to about 17.37 Mb/s. The MCR plus equal share of the six ABR VCs having an MCR of 1 Mb/s converged to about 13.37 Mb/s. The actual transmission rate of the four ABR VCs having an MCR of 5 Mb/s converged to about 17.37. The actual transmission rate of the six ABR VCs having an MCR of 1 Mb/s converged to about 13.37 Mb/s. The total ABR queue length stabilized around 1K cells. A very high link utilization (over 95%) was achieved.

The fourth experiment used fifteen ABR VCs each having an MCR of 1 Mb/s and ten ABR VCs each having an MCR of 5 Mb/s. All VCs had an ICR of 5 Mb/s and a PCR of 149.76 Mb/s. All VCs started transmitting at time zero with a rate equal to the ICR. The results of this experiment were that the MCR plus equal share of the fifteen ABR VCs each having an MCR of 1 Mb/s converged to about 4.39 Mb/s. The MCR plus equal share of the ten ABR VCs each having an MCR of 5 Mb/s converged to about 8.39 Mb/s. The actual transmission rate of the fifteen ABR VCs each having an MCR of 1 Mb/s converged to about 4.39 Mb/s. The actual transmission rate of the ten ABR VCs each having an MCR of 5 Mb/s converged to about 8.39 Mb/s. The total ABR queue length stabilized around 300 cells, and a very high link utilization (near 99%) was achieved.

#### Experiments 5 and 6

Two experiments were performed with ABR VCs having different MCRs (the same conditions as experiments three and four) and with bursty VBR background traffic: one with ten ABR VCs, and one with twenty-five ABR VCs. In these experiments a deterministic ON/OFF VBR source with 169 msec ON and 339 msec OFF periods was used. During ON period the VBR source

transmitted at 50 Mb/s and was idle during OFF period. The presence of bursty high priority traffic has been generally shown to cause fairness problems for ABR VCs. Since VBR traffic is scheduled with high priority, the presence of ON/OFF VBR traffic results in changing the drain rate of the ABR VCs. When the VBR source is ON, the bandwidth available for ABR VCs (above MCR) decreases. As a result, if these VCs were transmitting at a rate above MCR the queue occupancy of ABR VCs should increase.

The fifth experiment showed that when the VBR source was ON, the actual transmission rate of the four ABR VCs having an MCR of 5 Mb/s dropped from 17.37 Mb/s to 12.37 Mb/s. Similarly, the actual transmission rate of the six ABR VCs having an MCR of 1 Mb/s dropped from 13.37 Mb/s to 8.37 Mb/s. The ACR of the ABR sources showed that the algorithm does keep track of the changes in the ER. The queue occupancy of ABR VCs fluctuated with a frequency equal to that of the VBR source.

The sixth experiment showed that when the VBR source was ON, the actual transmission rate of the fifteen ABR VCs having an MCR of 1 Mb/s dropped from 4.39 Mb/s to 2.39 Mb/s. Similarly, the ten ABR VCs having an MCR of 5 Mb/s dropped from 8.39 Mb/s to 6.39 Mb/s. The ACR of the ABR sources showed that the algorithm does keep track of changes in the ER. The queue occupancy of ABR VCs fluctuated with a frequency equal to that of the VBR source.

#### Experiment 7

One experiment was performed with twenty-five ABR VCs having different MCRs (the same conditions as experiment four) and with bursty UBR background traffic. In this experiment a deterministic ON/OFF UBR source with 85 msec ON and 763 msec OFF periods was used. During the ON period the UBR source transmitted at 10 Mb/s and was idle during the OFF periods.

As mentioned above, a potential problem with dividing the available bandwidth equally on the basis of established connections (contending for the round-robin scheduler part) is that such a scheme can lead to under-utilization of the link when ON/OFF (bursty) sources are present. In order to avoid this, the algorithm keeps track of the active connections. The algorithm is self-correcting in the sense that as the queue size of a VC falls below the global queue threshold the available bandwidth is divided among the number of active connections rather than the number of established connections. This results in maintaining an overall high link utilization. For example, if the link is underutilized the per-VC queues will be drained at a higher rate. This results in the per-VC queues falling below the global queue threshold. When this happens, the ER calculations are done based on the number of active rather than the number of established connections which are contending for the available bandwidth. If some connections are OFF, the end result is that the algorithm allocates higher ER to the active VCs and hence maintains high link utilization even when ON/OFF connections are present.

The results of experiment seven showed that when the UBR source was OFF, the available bandwidth was divided among 25 ABR VCs connections. When the UBR source was ON, the available bandwidth was divided among 26 connections (i.e., 25 ABR VCs + 1 UBR VC). The queue length of ABR VCs fluctuated with a frequency equal to that of the ON/OFF UBR source. High link utilization was achieved.

There have been described and illustrated herein several embodiments of an algorithm for determining the explicit rate for an ABR connection in an ATM switch. While particular embodiments of the invention have been described, it is not intended that the invention be limited thereto, as it is intended that the invention be as broad in scope as the art will allow and that the specification be read likewise. It will therefore be appreciated by those skilled in the art that yet



other modifications could be made to the provided invention without deviating from its spirit and scope as so claimed.

## Claims:

1. A method for determining the explicit rate for an ABR connection in an ATM switch which serves guaranteed traffic and best efforts connections, comprising:

a) determining the available bandwidth in the switch after guaranteed traffic is served;

b) determining whether the ABR connection is congested;

c) if the ABR connection is congested, determining a fair share of available bandwidth as a function of the number of established best efforts connections in the switch;

d) if the ABR connection is not congested, determining a fair share of available bandwidth as a function of the number of active best efforts connections in the switch;

e) determining a new explicit rate as a function of the fair share; and

f) replacing the explicit rate for the connection with the new explicit rate if the explicit rate is larger than the new explicit rate.

2. A method according to claim 1, wherein:

said step of determining the available bandwidth is performed as a function of the total bandwidth of the switch and the bandwidth used by the guaranteed traffic.

3. A method according to claim 2, wherein:

the switch includes an egress side scheduler having a non-work-conserving part serving the guaranteed traffic and a round-robin part serving the ABR traffic; and

said step of determining the available bandwidth is performed according to the following equation

$$A_{bw} = \left(1.0 - \frac{U}{W}\right) \cdot LinkRate$$

where  $A_{bw}$  is the available bandwidth,  $U$  is the number of cells served by the non-work-conserving part during an interval of  $W$  cell slots and  $LinkRate$  is the total bandwidth of the switch.

4. A method according to claim 3, wherein:  
the interval  $W$  is chosen based on the *LinkRate*.
5. A method according to claim 4 wherein:  
for a *LinkRate* of approximately 149.76 Mb/s,  $W$  is approximately 1024 cells.
6. A method according to claim 1, wherein:  
each ABR VC in the switch has its own queue; and  
said step of determining whether the ABR connection is congested includes comparing the number of cells in the ABR connection queue with a threshold number.
7. A method according to claim 1, wherein:  
said step of determining a fair share of available bandwidth as a function of the number of established best efforts connections in the switch is performed according to the following equation

$$F_{\text{share}} = \frac{A_{\text{bw}}}{N_{\text{est}}}$$

where  $F_{\text{share}}$  is the fair share,  $A_{\text{bw}}$  is the available bandwidth, and  $N_{\text{est}}$  is the number of established best efforts connections in the switch.

8. A method according to claim 1, wherein:  
said step of determining a fair share of available bandwidth as a function of the number of active ABR connections in the switch is performed according to the following equation

$$F_{\text{share}} = \frac{A_{\text{bw}}}{N_{\text{act}}}$$

where  $F_{share}$  is the fair share,  $A_{bw}$  is the available bandwidth, and  $N_{act}$  is the number of best efforts ABR connections in the switch.

9. A method according to claim 1, wherein:

if the ABR connection is congested, the fair share is reduced by a damping factor before determining the new explicit rate.

10. A method according to claim 9, wherein

each ABR VC in the switch has its own queue;

the damping factor is a function of the number of cells in the ABR connection queue.

11. A method according to claim 10, wherein:

the damping factor is expressed as

$$\text{Max}\left(0, \left(1.0 - \frac{Q_{vc} - Q_{thresh}}{Q_{max}}\right)\right)$$

where  $Q_{vc}$  is the number of cells in the ABR connection queue,  $Q_{thresh}$  is a threshold number used to determine whether the ABR connection is congested, and  $Q_{max}$  is the maximum length of the ABR connection queue.

12. A method according to claim 8, wherein:

$N_{act}$  is a weighted average of the number of active best efforts connections.

13. A method according to claim 12, wherein:

$N_{act}$  is computed according to the following formula

$$N_{act} = (1 - \alpha)N_{old} + \alpha N_{new}$$

where  $\alpha$  is a weighting factor,  $N_{old}$  is the last measure of the number of active best efforts connections, and  $N_{new}$  is the present measure of the number of active best efforts connections.

14. An apparatus for determining the explicit rate for an ABR connection in an ATM switch serving guaranteed traffic and best efforts connections, comprising:

a) means for determining the available bandwidth in the switch after guaranteed traffic is served;

b) means for determining whether the ABR connection is congested;

c) means for determining a fair share of available bandwidth as a function of the number of established best efforts connections in the switch if the ABR connection is congested;

d) means for determining a fair share of available bandwidth as a function of the number of active best efforts connections in the switch if the ABR connection is not congested;

e) means for determining a new explicit rate as a function of the fair share; and

f) means for replacing the explicit rate for the connection with the new explicit rate if the explicit rate is larger than the new explicit rate.

15. An apparatus according to claim 14, wherein:

said means for determining the available bandwidth utilizes a function of the total bandwidth of the switch and the bandwidth used by the guaranteed traffic.

16. An apparatus according to claim 15, wherein:

the switch includes an egress side scheduler having a non-work-conserving part serving the guaranteed traffic and a round-robin part serving the ABR traffic; and

said means for determining the available bandwidth utilizes the following equation

$$A_{bw} = \left(1.0 - \frac{U}{W}\right) \cdot \text{LinkRate}$$

where  $A_{bw}$  is the available bandwidth,  $U$  is the number of cells served by the non-work-conserving part during an interval of  $W$  cell slots and  $\text{LinkRate}$  is the total bandwidth of the switch.

17. An apparatus according to claim 16, wherein:  
the interval  $W$  is chosen based on the  $\text{LinkRate}$ .

18. An apparatus according to claim 17 wherein:  
for a  $\text{LinkRate}$  of approximately 149.76 Mb/s,  $W$  is approximately 1024 cells.

19. An apparatus according to claim 14, wherein:  
each ABR VC in the switch has its own queue; and  
said means for determining whether the ABR connection is congested compares the number of cells in the ABR connection queue with a threshold number.

20. An apparatus according to claim 14, wherein:  
said means for determining a fair share of available bandwidth as a function of the number of established best efforts connections in the switch utilizes the following equation

$$F_{\text{share}} = \frac{A_{bw}}{N_{\text{est}}}$$

where  $F_{\text{share}}$  is the fair share,  $A_{bw}$  is the available bandwidth, and  $N_{\text{est}}$  is the number of established best efforts connections in the switch.

21. An apparatus according to claim 14, wherein:  
 said means for determining a fair share of available bandwidth as a function of the number of active best efforts connections in the switch utilizes the following equation

$$F_{\text{share}} = \frac{A_{\text{bw}}}{N_{\text{act}}}$$

where  $F_{\text{share}}$  is the fair share,  $A_{\text{bw}}$  is the available bandwidth, and  $N_{\text{act}}$  is the number of active best efforts connections in the switch.

22. An apparatus according to claim 14, further comprising:  
 g) means for reducing the fair share by a damping factor if the ABR connection is congested.
23. An apparatus according to claim 22, wherein  
 each ABR VC in the switch has its own queue;  
 the damping factor is a function of the number of cells in the ABR connection queue.
24. An apparatus according to claim 23, wherein:  
 the damping factor is expressed as

$$\text{Max}\left(0, \left(1.0 - \frac{Q_{\text{vc}} - Q_{\text{thresh}}}{Q_{\text{max}}}\right)\right)$$

where  $Q_{\text{vc}}$  is the number of cells in the ABR connection queue,  $Q_{\text{thresh}}$  is a threshold number used to determine whether the ABR connection is congested, and  $Q_{\text{max}}$  is the maximum length of the ABR connection queue.

25. An apparatus according to claim 21, wherein:  
 $N_{\text{act}}$  is a weighted average of the number of active best efforts connections.

26. An apparatus according to claim 25, wherein:

$N_{act}$  is computed according to the following formula

$$N_{act} = (1 - \alpha)N_{old} + \alpha N_{new}$$

where  $\alpha$  is a weighting factor,  $N_{old}$  is the last measure of the number of active best efforts connections, and  $N_{new}$  is the present measure of the number of active best efforts connections.

27. A method for determining the explicit rate for an ABR connection in an ATM switch serving guaranteed traffic and best efforts connections, comprising:

- a) determining the available bandwidth in the switch after guaranteed traffic is served;
- b) determining whether the ABR connection is congested;
- c) determining a fair share of available bandwidth;
- d) if the ABR connection is congested, reducing the fair share of available bandwidth by a damping factor;
- e) if the ABR connection is congested, determining a new explicit rate as a function of the reduced fair share; and
- f) replacing the explicit rate for the connection with the new explicit rate if the explicit rate is larger than the new explicit rate.

28. A method for determining the explicit rate for an ABR connection in an ATM switch serving guaranteed traffic and best efforts connections, comprising:

- a) determining the available bandwidth in the switch after guaranteed traffic is served;
- b) determining whether the ABR connection is congested;
- c) if the ABR connection is not congested, determining the number of active best efforts connections in the switch at time  $t_1$  and time  $t_2$ ;
- d) if the ABR connection is not congested, determining a fair share of available bandwidth as a function of a weighted average of the number of active best efforts connections in the switch



at time  $t_1$  and time  $t_2$ ;

e) determining a new explicit rate as a function of the fair share; and

f) replacing the explicit rate for the connection with the new explicit rate if the explicit rate is larger than the new explicit rate.

1/2

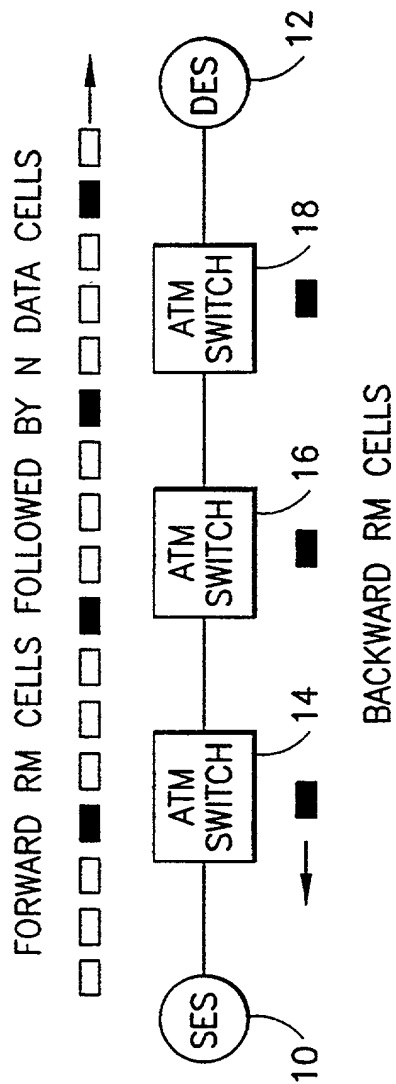


FIG.1

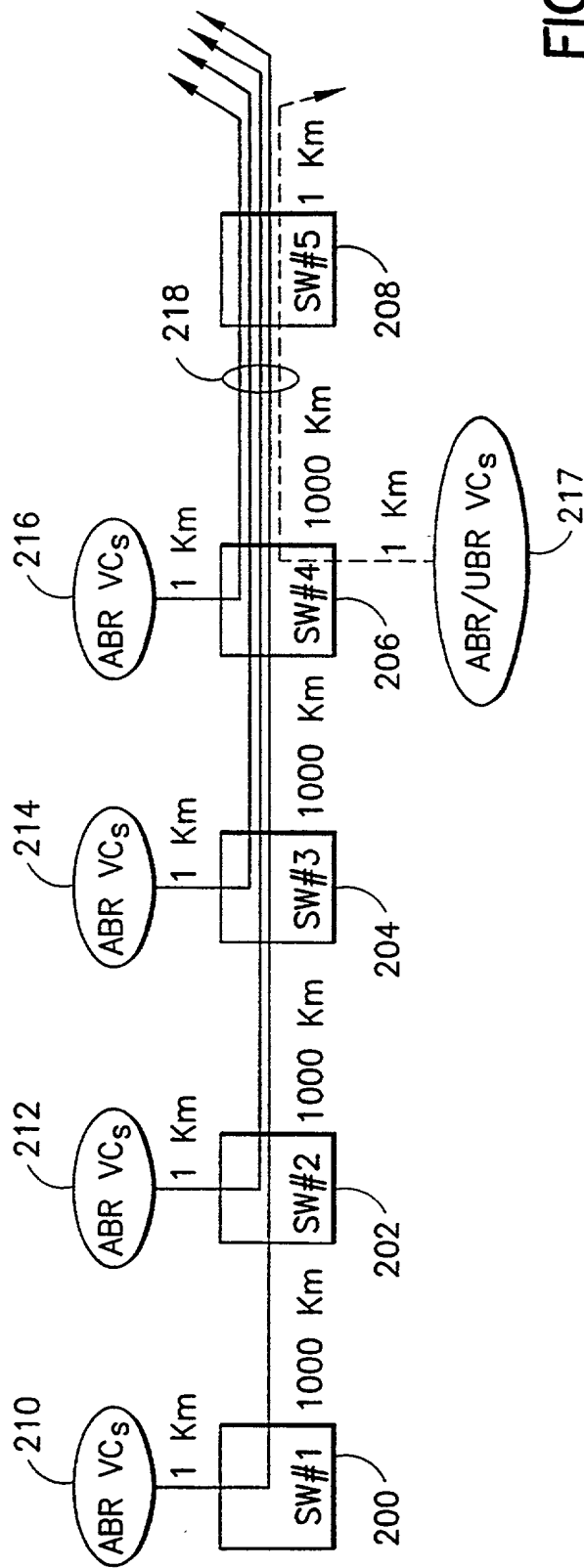


FIG.3

2/2

09/744323

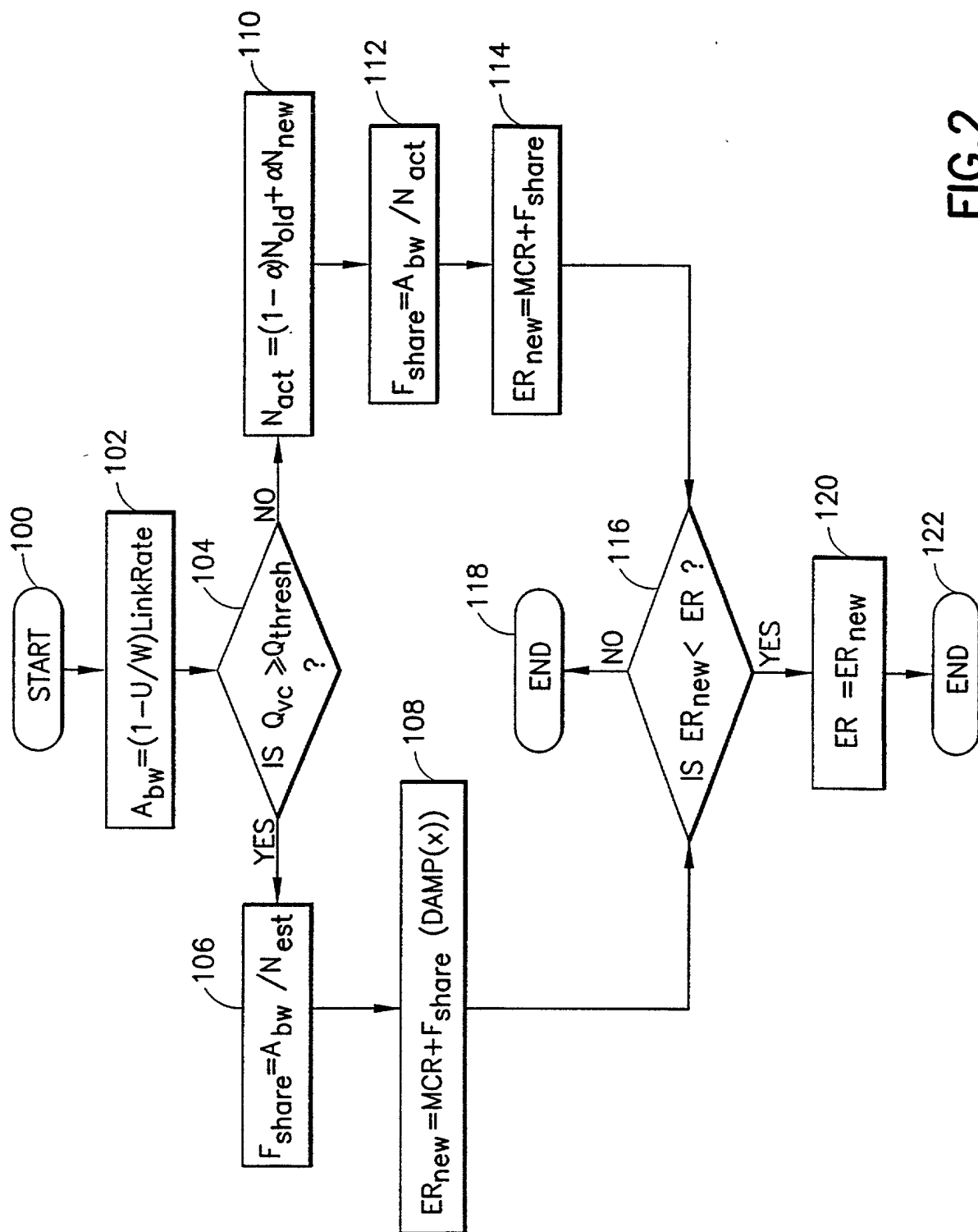


FIG. 2

Attorney Docket No.: GDC-129

**DECLARATION FOR PATENT APPLICATION AND POWER OF ATTORNEY**

As below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed for and for which a patent is sought on the invention entitled

**EXPLICIT RATE ABR ALGORITHM FOR USE IN AN ATM SWITCH,**

the specification of which

☐ is attached hereto.

☒ was filed on: January 19, 2001

as application Serial Number: 09/744,323

and was amended on (if applicable):

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, Section 1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Serial No.	Country	Filing Date (D/M/YR)	Priority Claimed?
1. PCT/US99/16810	PCT/US	22/07/99	<input checked="" type="checkbox"/> YES <input type="checkbox"/> NO
2.			<input type="checkbox"/> YES <input type="checkbox"/> NO

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Section 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

Application Ser. No.	Filing Date	Status (patented, pending, abandoned)
1.		
2.		
3.		

As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

David P. Gordon (29,996)  
David S. Jacobson (39,235)  
Thomas A. Gallagher (31,358)

Address all telephone calls to: David P. Gordon at (203) 329-1160  
Address all correspondence to: David P. Gordon, Esq.  
65 Woods End Road  
Stamford, Connecticut 06905  
U.S.A.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

#### SOLE OR FIRST INVENTOR

Signature: Iftexhar Hussain Date 01/22/2001

Full Name: Iftexhar Hussain

Residence: 2613 SOUTH PARK LANE, SANTA CLARA, CA 95051  
~~3770 Flora Vista Avenue, Apartment 1801, Santa Clara, CA 95051~~

Citizenship: Pakistan

P.O. Address: Same as address

#### SECOND INVENTOR

Signature: \_\_\_\_\_ Date \_\_\_\_\_

Full Name: Kuldip Bains

Residence: 665 Park Road Extension, Middlebury, CT 06762

Citizenship: GB

P.O. Address: Same as address

Application Ser. No.	Filing Date	Status (patented, pending, abandoned)
----------------------	-------------	--

1.  
2.  
3.

As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

3  
David P. Gordon (29,996)  
David S. Jacobson (39,235)  
Thomas A. Gallagher (31,358)

Address all telephone calls to: David P. Gordon at (203) 329-1160  
Address all correspondence to: David P. Gordon, Esq.  
65 Woods End Road  
Stamford, Connecticut 06905  
U.S.A.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

SOLE OR FIRST INVENTOR

Signature: \_\_\_\_\_ Date \_\_\_\_\_

Full Name: Iftexhar Hussain

Residence: 2613 South Park Lane, Santa Clara, CA 95051

Citizenship: Pakistan

P.O. Address: Same as address

SECOND INVENTOR

Signature:  Date 5/15/01

2-00 Full Name: Kuldip Bains

Residence: 665 Park Road Extension, Middlebury, CT 06762 CT

Citizenship: GB

P.O. Address: Same as address